Topical review

Examining the evidence about psychological treatments for chronic pain:
Time for a paradigm shift?

Stephen Morley a,⇑, Amanda Williams b, Christopher Eccleston c

aUniversity of Leeds, UK
cUniversity College London, UK
cUniversity of Bath, UK

1. Introduction

Psychological interventions are now well established as a core part of modern pain practice. Randomized controlled trials (RCT) of psychological treatment for chronic pain were first performed in the 1970s, and considerable effort, skill, and sophistication have been applied to establishing the evidence of effectiveness [26]. Over the last 2 decades, we and others have conducted systematic reviews and meta-analyses of psychological interventions, predominantly cognitive behaviour therapy (CBT), for chronic pain [3,6,8,16,17,27,40]. Meta-analysis facilitates the emergence of underlying patterns in data by controlling the bias and error inherent in individual studies. These meta-analyses suggest that, overall, CBT has a beneficial average effect for a range of outcomes, principally disability, depression, and pain experience. Evidence of harm is not available. Although the headline effect is positive, current analyses are limited to average results, and conclusions are typically confined to the general.

The aim of this topical review is to consider the next steps in developing and evaluating psychological treatments for chronic pain. In this examination, we take as our unit of analysis the meta-analyses rather than individual RCTs. We briefly summarize the results of the meta-analyses, their benefits, and their limits, then offer directions for improving the next generation of studies. We argue that we have reached a critical point in the evolution of psychological interventions, and a paradigm shift is now needed in how we investigate treatment efficacy, effectiveness, and harm.

2. Evidence from meta-analysis

Our first meta-analysis [27] found CBT to be more effective (comparing groups after treatment) than no treatment/treatment as usual for outcomes of pain, pain experience, cognitive coping and appraisal, behavioral expression of pain, mood/affect, and social role functioning, with a median effect size ES(d) of 0.5. There was marginal evidence that CBT was superior to other active treatments. In later analyses, using improved methods, CBT remained superior to no treatment/treatment as usual, but the magnitude of d diminished to around 0.2 [8,40], similar to effect sizes in other meta-analyses for mixed chronic pain [33], low back pain [16,17], fibromyalgia [3,11,14], and arthritis [6]. The smaller recent effect sizes are likely as a result of improvements in the control of bias, such as a stricter criterion for entry on individual study sample size [30]. The methodological quality in the design of trials, including risk of bias, improved over time, but treatment quality did not (see Fig. 2 in [26]). This held after excluding many trials for inadequate/insufficient psychotherapeutic content. Despite having more and better-designed trials for meta-analysis, the picture remains unclear. Half of the comparisons found no effect of CBT and half found weak effect sizes of unknown clinical significance on pain, mood, disability and catastrophic thinking outcomes [40]. The clarity of any overall effect is muddied by three main sources: (1) sample, measure, and treatment heterogeneity, (2) unexplained variation of outcome reporting across studies, and (3) theoretical imprecision.

In summary, current meta-analyses indicate that psychological treatments are likely to be effective. However, as more studies are produced and added to meta-analyses, the results are effectively being diluted by the addition of poorly conceived, conducted, and reported trials. The signal-to-noise ratio is shrinking. Meta-analysis cannot control for poor primary studies or poor reporting [20]. Additional trials that are atheoretical, biased, single center, and enthusiasm driven will make summary and interpretation increasingly difficult. We suggest that without radical change, simply adding further trials will not help us to improve treatment effectiveness nor help answer research questions.

This is not a counsel of despair. We are not arguing for the abandonment of the program of developing and testing psychological treatments for chronic pain. Psychological therapies for the management of chronic pain offer worthwhile and promising treatments. Rather, we seek to build on the evidence base, but we suggest changing the methods of investigation [22,32]. If we do not improve the methods of primary investigation and learn from the trial developments in other areas of pain investigation, then we will achieve only confusion [32].
3. The complexity of chronic pain

Chronic pain is complexly maintained, and its resistance to treatment should not surprise us. In fact, it recently has been argued that analgesic failure is to be expected and should be a guiding treatment principle of all chronic pain treatment [21]. Patients present with disability established over many years. Many of the trials present treatments of brief duration, sometimes delivered by personnel trained just for that trial; good clinical outcomes are rather less likely from dilute/short treatments delivered by inexperienced staff to severely distressed patients. Perhaps we have underestimated the complexity of behavior change and the social and psychological influences that maintain disability in chronic pain patients [4]. Even if pain can be modulated by non-psychological methods, the complexity of psychological and behavioral adaptation means that the interruptive, interfering, and identity-distorting impacts of pain are no longer functionally related to the immediate experience of pain but are controlled by contextual behavioral, social, and cognitive factors that require therapeutic attention [10,18,23].

4. Heterogeneity in patient populations and treatment complexity

Many trials have mixed patient samples made up of patients with many and varied diagnoses. Critics have argued for greater specificity, and trials are emerging of single-diagnosis patient samples. Identifying groups of patients by disease or disorder may aid evidence-based translation to clinical practice, but psychological profiles are often orthogonal to medical diagnosis [9]. Diagnostic group membership is largely unhelpful in explaining patient pain behaviour or in guiding treatment decisions. Entry criteria for trials rarely specify levels of distress or disability, other than to exclude participants with intellectual limitations, overt symptoms of psychosis, or addiction. The distribution of scores on measures of depression and distress covers the full spectrum. A psychologically informed characterization of patients might facilitate better-targeted and more effective treatment, but the optimal method for patient group determination is still under debate [29,36,38]. Many of the trials entered into meta-analyses comprise pragmatic mixes of treatment content. The rationale for treatment choice and blend in multicomponent treatments is often unreported, and perhaps of more concern, there is often a disparity between the aims of treatment, the actual treatment content, and outcomes reported. This makes it difficult to discern which components of treatment contribute to specific changes. Component dismantling studies offer an illusion of identifying active ingredients but cannot achieve sufficient power to calculate the effects of each component on each outcome [12]. It is worth noting that within the general field of psychotherapy, it is difficult to obtain effect sizes significantly different from zero for the difference between treatment packages with and without a putatively effective component [1].

5. Outcome measures and measurement

Historically, psychological treatments have utilized outcomes with continuous measures that are expressed by Cohen’s ES($d$) and its variants. This has inherent practical limitations and implications. The between-group ES of 0.2, typical of recent meta-analyses, is unhelpful in communication with patients who want to know what their chances are of getting better; replying, ‘The average person in the treated group is at the 58th percentile (or 69th percentile when ($d$) = 0.5) of the untreated group,’ although accurate, is unhelpful [24]. Most trials report statistical rather than clinical significance. Rarely do trials report binary outcomes that are based on clinical significance criteria [24,29]. Consequently, we have no basis on which to estimate the number of successfully treated people. Pre–post treatment effect sizes are also modest (~0.4) and do not approach the value of around 1.5 desired by patients [31,34]. Finally, although diverse measures are reported in trials, it is not clear that they necessarily coincide with the interests of patients [2,35]. Nor are adverse effects or dropout rates adequately recorded [28]. This may reflect inadequate preparation for treatment—or worse, a failure of equipoise in which investigators simply do not expect failure.

6. The next steps

The treatment of chronic pain remains a challenge whether the primary aim is analgesic or rehabilitative. The success rates (and effect sizes) for psychological treatments are similar to those reported for pharmacological and surgical interventions [5,15]. What is different from the evidence base for nonpsychological interventions is that as more trials of psychological treatments are published, clarity becomes more, not less, elusive. Additional low-quality trials of minor variations in therapy will not solve the problems outlined here and will only add noise. A paradigm shift is necessary and long overdue. Three major advances are needed. We invite consideration and debate.

First, we encourage the design of theoretically coherent evaluations and trials. Precise testable models are needed that link specific treatment procedures with specific psychological changes leading to hypothesized specific outcomes. Therapy needs to be based on explicit theoretical models that guide choice of content, dose, timing and quality of treatment, and the choice of specific outcomes. There are examples of useful theory. The evidence for graded exposure has been developed in a clinical context using single case methodology to systematically replicate and test its component processes [37]. The theoretically sophisticated functional affective–motivational model of attention and pain has been extensively explored and tested in laboratory settings [19]. It is markedly different from the limited channel capacity model on which most clinical interventions for attention modulation are based [25] and is ripe for translation.

Second, we propose a standardizing of key features of methodological quality and bias. In particular, we should adopt the IMMPACT criteria on measurement domains and measurement tools [7]. We should judge small-sample-size RCTs (including those labeled ‘feasibility,’ ‘preliminary,’ and ‘exploratory’) as low quality [13] and encourage studies with adequate control over bias. We need clear articulation of the severity of patient complaint using a shared language of the extent of disability, distress, and alterable behaviour [39]. Critical to bias control will be an honest appraisal of the common lack of equipoise in psychological trials, and the requirement at the very least for the transparent reporting of therapist allegiance. Finally, trials that do not measure or report adverse events should be considered fundamentally flawed.

Third, alternatives to the randomized controlled trial should be considered. Rowbotham et al. [32] recently argued that to properly address the clinical effectiveness of an efficacious treatment, we need to invest in large-scale observational and translational studies. This is also relevant in psychological interventions, and in particular for the benchmarking of achievable outcomes in regular clinical practice. Single case methodologies are also infrequently used and have promise [37]. Advances in secondary data analysis are also relevant, and analysis at the patient level across trials may be particularly relevant for indirect comparisons between treatments. Responder analyses are also promising and should be explored. Both will rely on open access to data.
We argue that it is time for a radical change in how we design treatments and study their effectiveness. A paradigm shift is essential. We have gone as far as we can with the old models. The next generation of studies will need to raise the bar on quality. Studies that do not measure adverse events should be considered unethical. Studies that do not control therapist allegiance should be considered flawed. Studies that are small should be ignored. The next studies will better define homogenous samples of patients and will match key sample features to treatment content and outcome assessment. Access to data at an individual patient level will allow for responder analyses. And access to larger data sets on community studies will allow for effectiveness studies at scale. Future systematic reviews will also lead to improvement by being explicit and transparent about trial entry. Evidence will be graded with tools that recognize the influence of bias in individual trials, and innovative methods of portraying the results of comparative effectiveness studies to different stakeholders will improve their relevance.

Conflict of interest

The authors report no conflict of interest.

References